Informatica PowerExchange for PADB User's Guide

Version 3.0

Contents

Important

About ParAccel™, Inc.

ParAccel is the developer of ParAccel Analytic Database (PADB), the world's fastest, most cost-effective extensible platform for accelerating insight and empowering analytics-driven businesses. ParAccel enables organizations to tackle the most complex analytic challenges and glean ultra-fast deep insights from vast volumes of data. Data-driven businesses in financial services, retail, health care, government and more are taking advantage of ParAccel to tackle critical, time-sensitive questions outside the scope of conventional data warehouses and existing analytic tools.

PADB is a columnar-based, massively parallel processing (MPP) software only analytic database. Designed for businesses seeking data insight at the speed of now without costly investment in proprietary hardware solutions, ParAccel delivers the industry's highest performance against complex SQL workloads and the most challenging analytic queries. Extensible in-database analytics, true schema-neutral, load-and-go analysis and unmatched price-performance means an accelerated speed to information and ROI.

For more information please contact us at *info@paraccel.com* or 866-903-0335, or visit us at *http://www.paraccel.com*.

Notices

Disclaimer

This document is not to be construed as a promise by any participating company to develop, deliver, or market a product. ParAccel, Inc., makes no representations or warranties with respect to the contents of this document, and specifically disclaims any express or implied warranties of merchantability, usefulness, fitness for a particular purpose, or noninfringement. Further, ParAccel, Inc., reserves the right to revise this document and to make changes to its content, at any time, without obligation to notify any person or entity of such revisions or changes.

Trademarks

Because of the nature of the material, numerous hardware and software products are mentioned by their trade names in this publication. ParAccel, Inc. and the ParAccel, Inc. logo, the ParAccel Analytic Database, ParAccel AMIGO, and ParAccel MAVERICK are trademarks of ParAccel, Inc.

Other brand and product names mentioned in this document may be the trademarks or registered trademarks of their respective owners.

ODBC, Microsoft SQL Server, SQL Server, and Internet Explorer are either registered trademarks or trademarks of Microsoft Corporation in the United States and/or other countries.

Red Hat is a registered trademark of Red Hat, Inc. Linux is a registered trademark of Linus Torvalds in the U.S. and other countries. PostgreSQL is a trademark of PostgreSQL Development Group.

Java, JDK, and JDBC are trademarks or registered trademarks of Sun Microsystems, Inc. or its subsidiaries in the United States and other countries.

AMD and Opteron are registered trademarks of AMD Technology Solutions LLC.

Intel and Xeon are registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries.

Veritas is a trademark of Symantec Corporation or its affiliates in the U.S. and other countries.

Liferay is a registered trademark of Liferay, Inc.

JasperServer and iReport are trademarks or registered trademarks of JasperSoft Corporation in the U.S. and other countries.

MicroStrategy is a registered trademark of MicroStrategy Incorporated in the United States and certain other countries.

DataDirect is a trademark or registered trademark of Progress Software Corporation or one of its subsidiaries or affiliates in the U.S. and other countries.

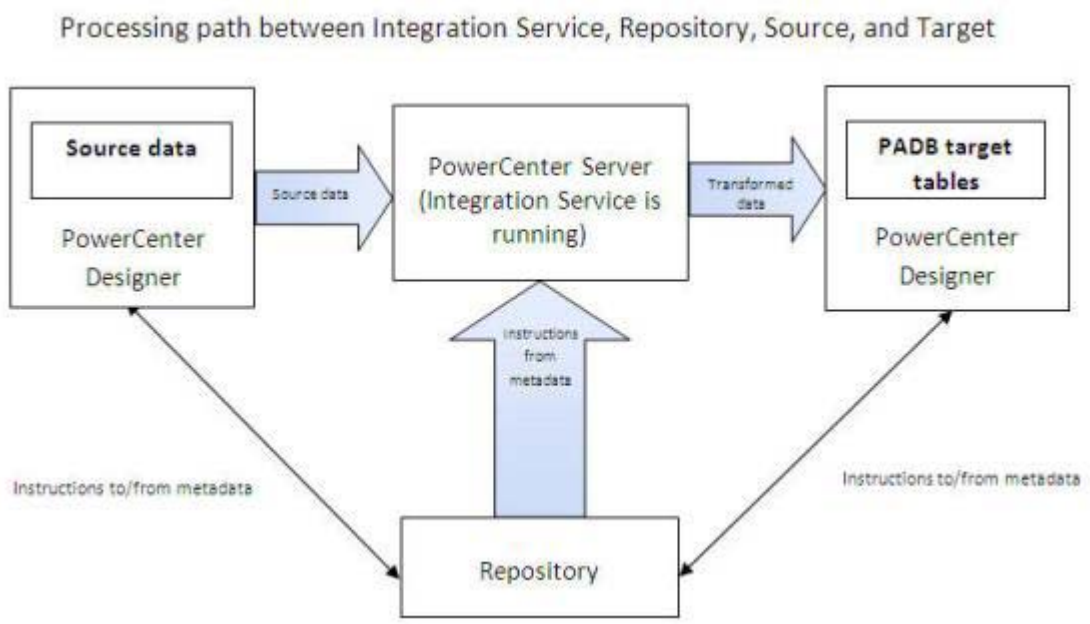
ANSI is a trademark of American National Standards Institute.

Python is a registered trademark of the Python Software Foundation.

Introduction to Informatica PowerExchange for PADB

Informatica PowerExchange for PADB integrates PowerCenter with PADB by using the COPY command to load data into PADB tables. You can use PADB tables as targets in the PowerCenter Target Designer and load data from any source supported in the PowerCenter Source Analyzer. The sources can include flat files, XML files, relational tables, COBOL files, and so on.

Before loading data into a PADB table, you must import both target definitions and source definitions. Target definitions are the metadata of the PADB tables into which data will be written. Source definitions are the metadata of the tables or files whose data you want to load into PADB. The PowerCenter repository stores the metadata of target and source definitions. The metadata describes the schema of the imported tables, such as column names and data types.



The whole process of loading data into PADB is as follows:

1. Importing source definitions
2. Importing target definitions
3. Creating mappings
4. Creating sessions
5. Creating workflows

For details about these tasks, see *Loading data into PADB tables*, as well as the online help and documentation available with the PowerCenter tools.

PADB connector

The PADB connector is a plug-in that enables you to use PADB as a target in the PowerCenter Target Designer and load data from any source supported in the PowerCenter Source Analyzer. The sources can include flat files, XML files, relational tables, COBOL files, and so on.

The installation of PowerExchange for PADB requires the following files.

File Name	Description
pmpadbclient.dll	This file is the client plug-in used to import the PADB target definitions. See *Client plug-in*.
padb_EN.res	This file contains a list of messages. PowerExchange for PADB uses different RES files for different operating systems. Contact the ParAccel technical support team to obtain the appropriate RES file. **Note:** Use the file name `padb_EN.res` in the Linux environment.
pmpadbwriter.dll	This plug-in is a DLL file that defines the functionality of PowerExchange for PADB. When you register the plug-in, you add the functionality of the plug-in to the repository. The writer plug-in is deployed on the PowerCenter Server, where the Integration Service is running. The main functionality of the writer plug-in is to execute the target definition created in the mapping and load data into the PADB tables. See *Writer plug-in*.
PADBPlugin.xml	This XML file is used to register the writer plug-in in the Informatica PowerCenter Administration Console. See *Repository plug-in*.
exchange_padb.chm	This file contains the online help.

Note:

1. You have to use different ODBC drivers for different operating systems to connect to PADB. Contact ParAccel to obtain the appropriate ODBC driver.
2. Concurrent execution of sessions on the same table is not supported since PADB does not support the concurrent execution of COPY commands on the same table. The connector does not support the concurrent workflow feature.
3. PowerExchange for PADB supports only Unicode data movement. The connector reads UCS-2 encoded data from the PowerCenter memory buffer and uses ODBC calls in UTF-8 to write to PADB.
4. PowerExchange for PADB does not overcome:
   - Any limitations or known issues in functionality or performance of the PADB COPY command.
   - Any limitations or known issues in functionality or performance of Informatica PowerCenter.

Client plug-in

The client plug-in *pmpadbclient.dll* is used to import PADB target definitions. The target definitions contain metadata of the PADB tables. The client plug-in creates the **Import from PADB** menu item in the **Targets** menu of the PowerCenter Target Designer and connects to PADB by using ODBC.

Writer plug-in

The writer plug-in is a DLL file (on Windows) or .so file (on Linux) that enhances the functionality of PowerExchange for PADB. The writer plug-in executes the target definition created in a mapping and loads data into the corresponding PADB tables. You add the functionality of the plug-in to the repository created in Informatica when you register the repository plug-in in the PowerCenter Administration Console. The writer plug-in is deployed on the Informatica PowerCenter server, where the Integration Service is running.

The writer plug-in performs the following tasks:

1. Connects to PADB.
2. Sends a PADB COPY command using the STDIN option.
3. Reads from the Informatica buffers.
4. Converts the data types for the load.
5. Transmits the converted data to the running COPY command.
6. Writes error messages in the PowerCenter session log. These error messages are generated by PADB during data loading.

The writer plug-in also supports the following features:

• Localization of limited Informatica session log messages
• The Session-on-a-grid feature of PowerCenter

PowerExchange for PADB uses different writer plug-in files for different operating systems, as shown in the following table.

Platform	Writer File
Windows 2003 Server 32-bit	pmpadbwriter.dll
Windows 2003 Server 64-bit	pmpadbwriter.dll
Red Hat Linux 64-bit	pmpadbwriter.so

Contact the ParAccel technical support team to obtain the appropriate writer file.

Repository plug-in

The `PADBPlugin.xml` file is an XML file that is used to register the repository plug-in in the PowerCenter Administration Console. It adds the functionality of the writer plug-in to PowerCenter. The XML file contains the following information:

Data	Description
Plug-in ID	Identifies the connector.
Data type support for ParAccel	Indicates support for ODBC data types.
List of writer library names for supported platforms	Indicates the version of the writer plug-in file to be used in different operating systems.
Connection type	Indicates that the PADB database connection is a relational connection.
Connection and session-level attributes	Indicates the attributes used by the writer plug-in.

Installing the connector plug-in

PowerExchange for PADB allows you to access the PADB metadata stored in the PowerCenter repository and run sessions against PADB. This section gives you information about installing and configuring PowerExchange for PADB on Microsoft Windows Server 2003 and Red Hat Linux 4.0 64-bit platforms.

You can install PowerExchange for PADB in the following environments:

- Microsoft Windows Server 2003 with 64-bit or 32-bit version of Informatica
- Red Hat Linux 4.0 64-bit version with 64-bit version of Informatica

There are two scenarios for installing PowerExchange for PADB:

- The Informatica server is installed on a Microsoft Windows Server 2003 machine and the Informatica client is installed on a Windows XP machine.
- The Informatica server is installed on a Red Hat Linux 4.0 machine and the Informatica client is installed on a Windows XP machine.

The installation and configuration process in both environments includes the following steps:

1. Installing prerequisites
2. Copying the necessary files
3. Registering the client plug-in in the Windows environment
4. Registering the repository writer plug-in
5. Establishing an ODBC connection

Installation prerequisites

PowerExchange 8.6.1.0.0 for ParAccel supports PowerCenter 8.6.1 and PADB 2.0.1x (PADB 2.5x recommended).

The PowerExchange for ParAccel client plug-in supports Power Center 8.6.1 on Windows. The writer plug-in is supported on the systems listed in the following table:

PowerCenter Integration Service OS	Version	Chip Set	Binary Type
Red Hat Linux	4	x86	64-bit
Microsoft Windows Server	2003	x64	64-bit
Microsoft Windows Server	2003	x86	32-bit

Note: If you will be installing the client plug-in on Windows XP or 2K3, you will need to have the Visual Studio 2008 run time already installed before installing the plug-in.

Installing the connector on Windows

PowerExchange for PADB allows you to access the PADB metadata stored in the PowerCenter repository and run sessions against PADB. This section assumes that Informatica PowerCenter is installed with SQL Server 2005.

Note: If you will be installing the client plug-in on Windows XP or 2K3, you will need to have the Visual Studio 2008 run time already installed before installing the plug-in.

If you are installing the connector on a machine where only the Informatica client is installed, follow these steps:

1. Copy each of the following files to the appropriate directory:
   - pmpadbclient.dll to `<PowerCenter 8.6.1 Installation Directory>\Client\Bin`
   - padb_EN.res to `<PowerCenter 8.6.1 Installation Directory>\Client\Bin`
   - exchange_padb.chm to `<PowerCenter 8.6.1 Installation Directory>\Client\Bin\Help\En`
2. Register the client plug-in in the Windows Registry Editor.
3. Establish and configure ODBC connectivity.

If you are installing the connector on a machine where only the Informatica server is installed, follow these steps:

1. Copy each of the following files to the appropriate directory:
   - padb_EN.res to `<PowerCenter 8.6.1 Installation Directory>\Client\Bin`
   - pmpadbwriter.dll to `<PowerCenter 8.6.1 Installation Directory>\Server\Bin`
   - PADBPlugin.xml a location accessible from your web browser
2. Register the repository plug-in.
3. Establish and configure ODBC connectivity.

If you are installing the connector on a machine where the Informatica client and server are both installed, follow these steps:

1. Copy each of the following files to the directory given below:
   - pmpadbclient.dll to `<PowerCenter 8.6.1 Installation Directory>\Client\Bin`
   - padb_EN.res to `<PowerCenter 8.6.1 Installation Directory>\Client\Bin`
   - pmpadbwriter.dll to `<PowerCenter 8.6.1 Installation Directory>\Server\Bin`
   - exchange_padb.chm to `<PowerCenter 8.6.1 Installation Directory>\Client\Bin\Help\En`
   - PADBPlugin.xml to a location accessible from your web browser
2. Register the client plug-in in the Windows Registry Editor.
3. Register the repository writer plug-in.
4. Establish ODBC connectivity.

Note: If you are using the 64-bit version of Windows 2003 Server, obtain the corresponding writer plug-in file from the ParAccel technical support team.

Registering the repository plug-in

Add the functionality of the plug-in to the repository created in Informatica when you register the repository plug-in in the PowerCenter Administration Console. You must have at least local administrator privileges to register the plug-in.

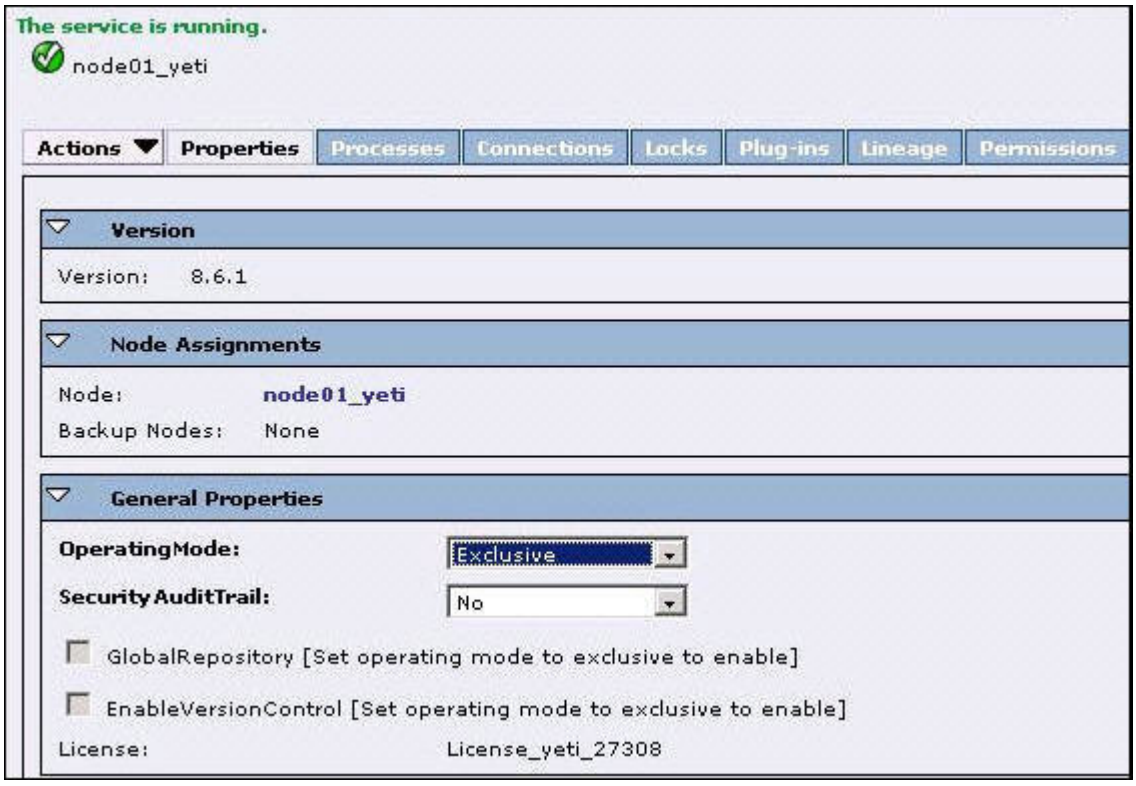
Use the `PADBPlugin.xml` file to register the writer plug-in in the PowerCenter repository. Change the repository mode to `Exclusive` mode before registering the repository plug-in file. After registering the plug-in, change the repository mode back to `Normal`.

Perform this procedure exactly once per repository, but do not repeat the procedure for each configured client. Clients should not attempt to use the plugin in the repository before the repository is registered.

General instructions for registering and un-registering a plug-in are located in the *PowerCenter Administrator Guide*.

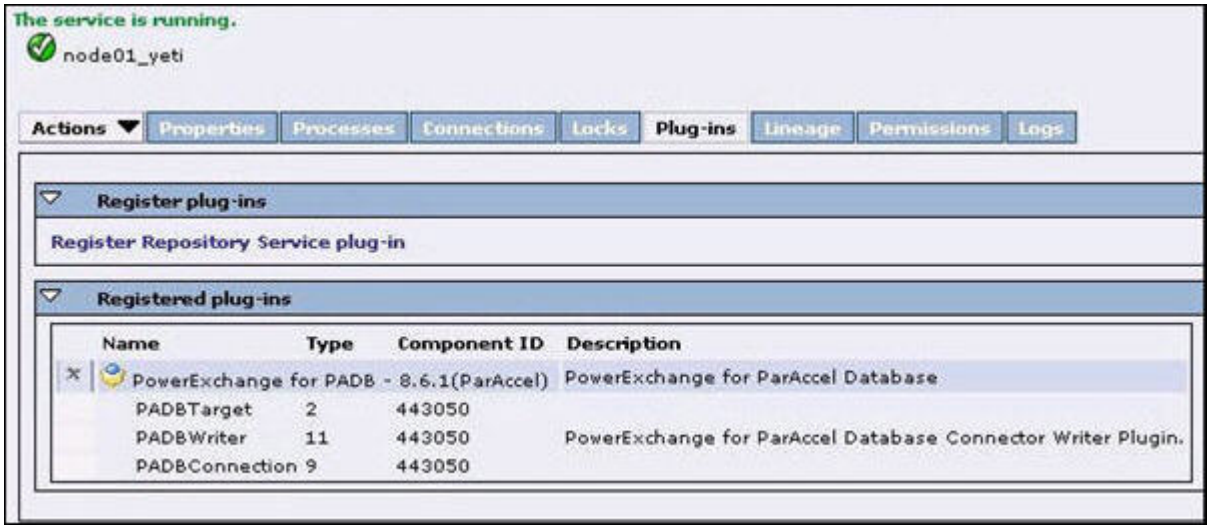
To register the plug-in:

1. Log into the PowerCenter Administration Console.
2. Change the mode of the repository to `Exclusive`.

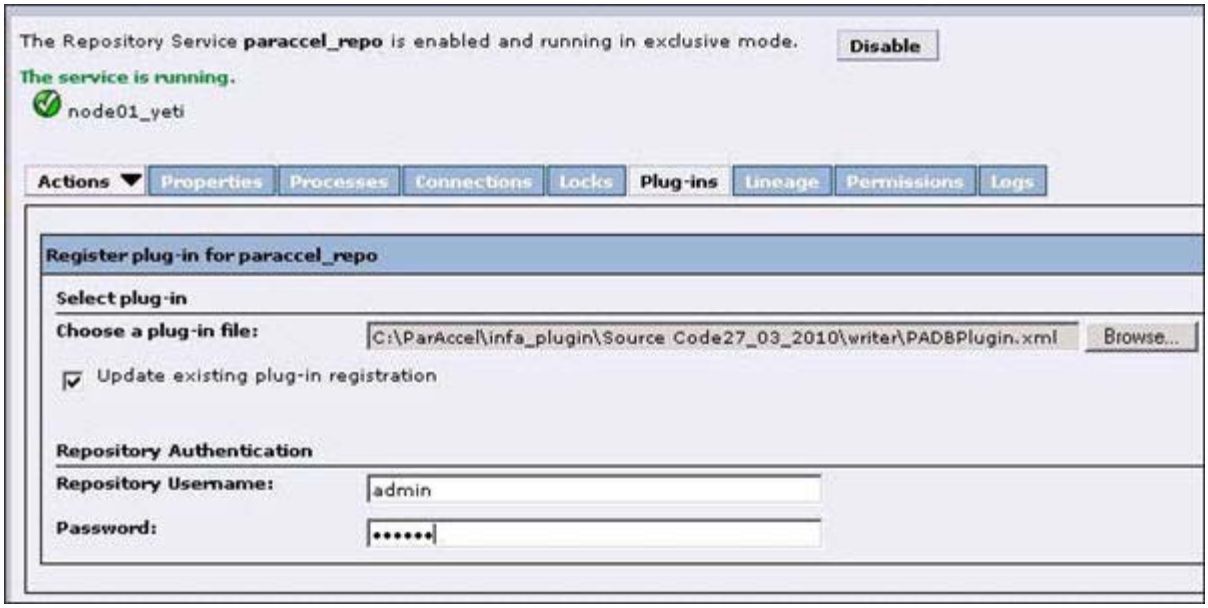


   Once the repository is switched to `Exclusive` mode, other users will not be able to access it.
3. Click the plug-ins tab to register the plug-in:

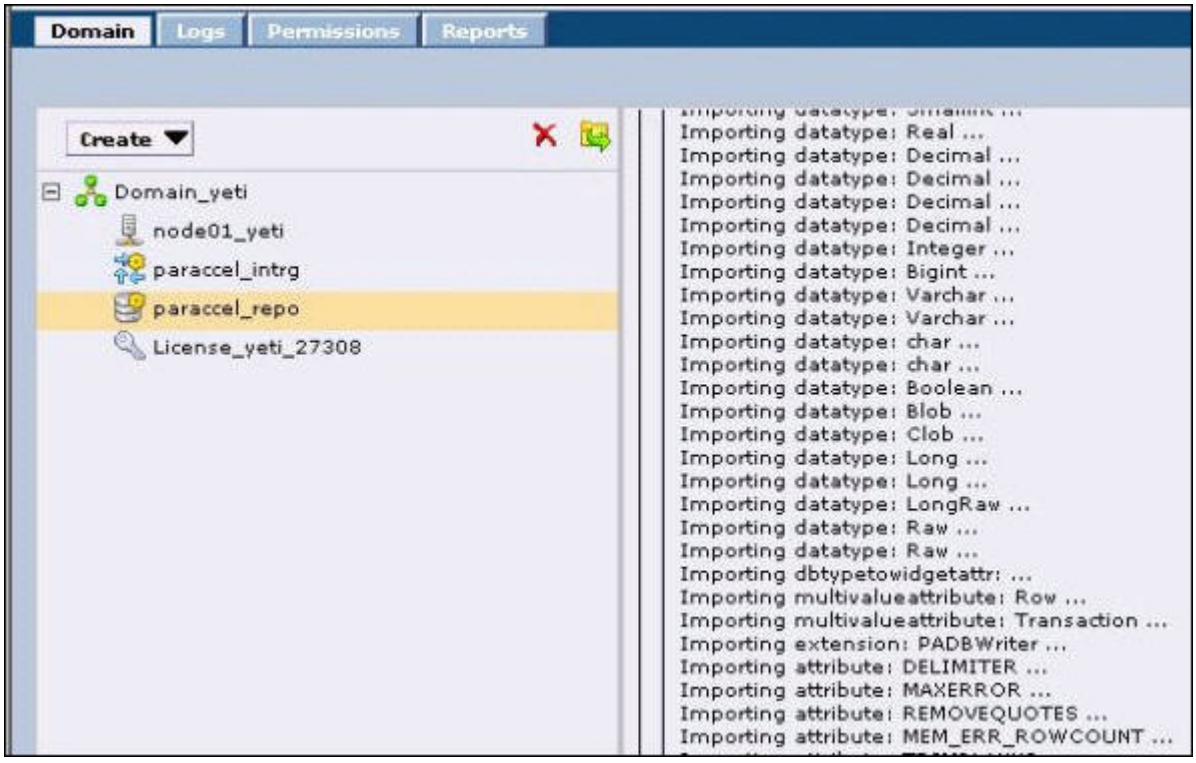
4. Wait a few seconds for a response. The plug-in name then appears in the list of registered plug-ins.



5. Click the **Register Repository Service** plug-in link and enter the information for the plug-in.



6. Click **Ok**.
   The XML file for the plug-in is applied.



7. Change the mode of the repository back to `normal`.

   The repository service is automatically re-started when you change the operating mode back to normal.

   For more information about installing the plug-in using the Administration Console, see the *PowerCenter Administrator Guide*.

   Note: You can also use the pmrep RegisterPlugin command to register the plug-in. For more information about this command, see the *PowerCenter Command Reference*.

Registering the client plug-in

The client plug-in file is used to import PADB target definitions. The plug-in creates the **Import from PADB** menu item in the **Targets** menu of the PowerCenter Target Designer. It connects to the PADB system by using ODBC.

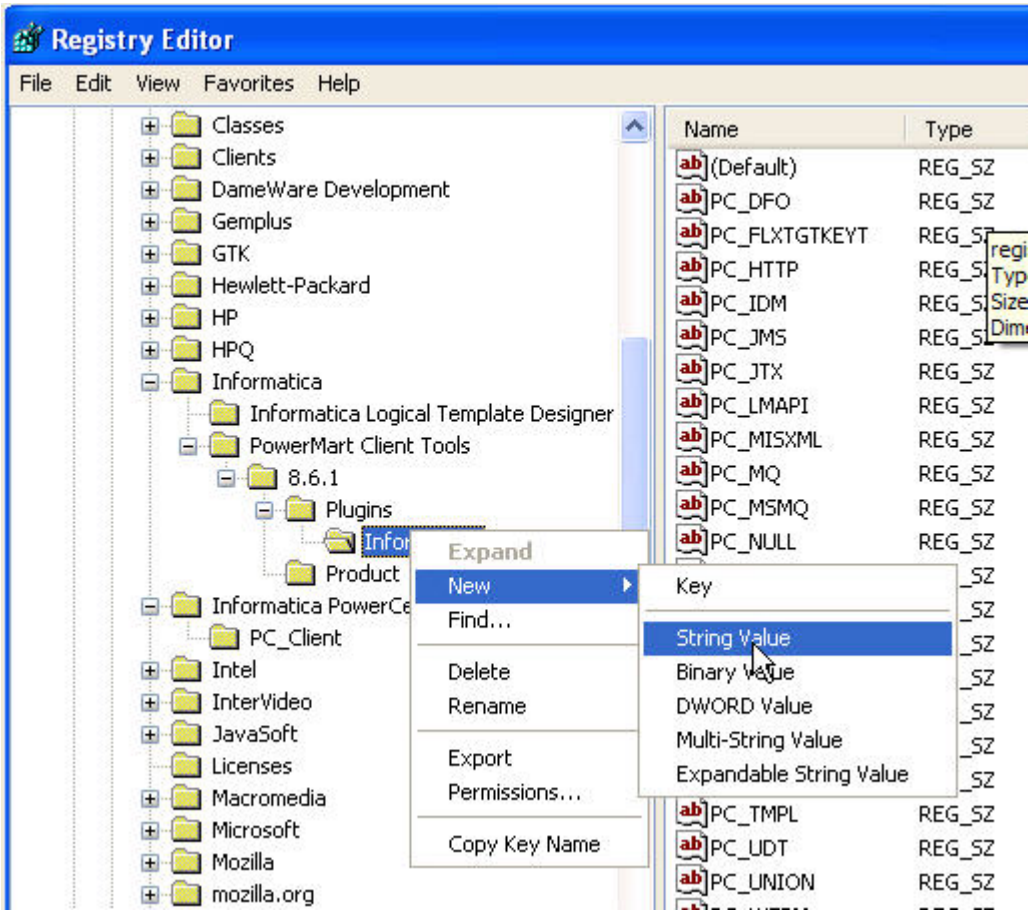
You have to register the client plug-in file, `pmpadbclient.dll`, in the Windows registry. This file is required to create or import a target definition in PowerCenter Designer.

To register the client plug-in:

1. Open the **Registry Editor** in Microsoft Windows Server 2003.
2. Open the following location in the **Registry Editor**:

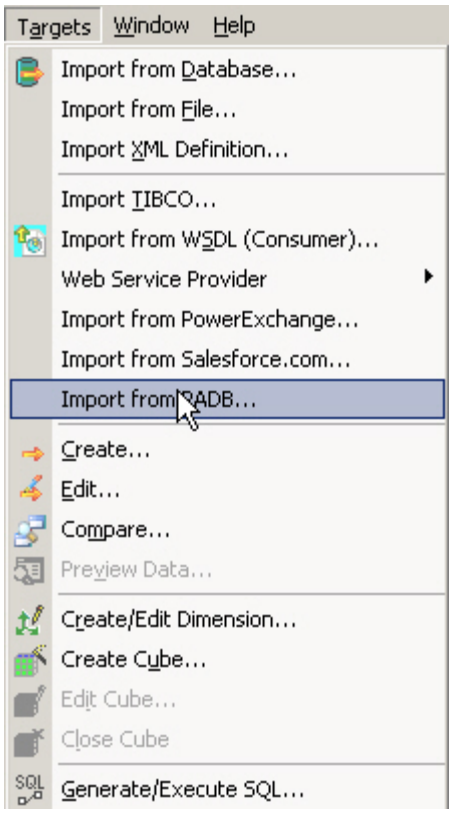
```
HKEY_LOCAL_MACHINE\SOFTWARE\Informatica\PowerMart
Client Tools\8.6.1\Plug-ins\Informatica
```

3. Right-click **Informatica** and create a new string `PC_PWXPADB`.



4. Add the value `pmpadbclient.dll` to `PC_PWXPADB`.
5. Close the **Registry Editor**.
6. Start Informatica Services.

You can now see the **Import from PADB** menu item in the **Targets** menu of the PowerCenter Designer.



Note: For more information about adding registry values in the Registry Editor, see the Registry Editor Help for Microsoft Windows Server 2003.

Caution: Incorrectly editing the registry may severely damage your system. Back up any important data on your computer before making changes to the registry.

Registering the writer plug-in

The writer plug-in executes the target definition created in a mapping and loads data into PADB tables. The writer plug-in is a DLL file that defines the functionality of PowerExchange for ParAccel. You add the functionality of the plug-in to the repository created in Informatica when you register the repository plug-in in PowerCenter Administration Console. The writer plug-in is copied on the PowerCenter Server, where the Integration Service is running.

Establishing an ODBC connection

PowerCenter uses Open Database Connectivity (ODBC) to connect to ParAccel and import ParAccel metadata. An ODBC data source contains information to access the database.

To establish an ODBC connection to PADB, create an ODBC data source on the PowerCenter client machine where you want to access ParAccel metadata. The client machine is where you run the PowerCenter Designer, Workflow Manager, Workflow Monitor, and Repository Manager. Use the ODBC Data Source Administrator and the ODBC driver provided by ParAccel to create and configure a data source.

For more information on how to establish a ParAccel ODBC connection, see the PADB *Administrator's Guide and SQL Reference*.

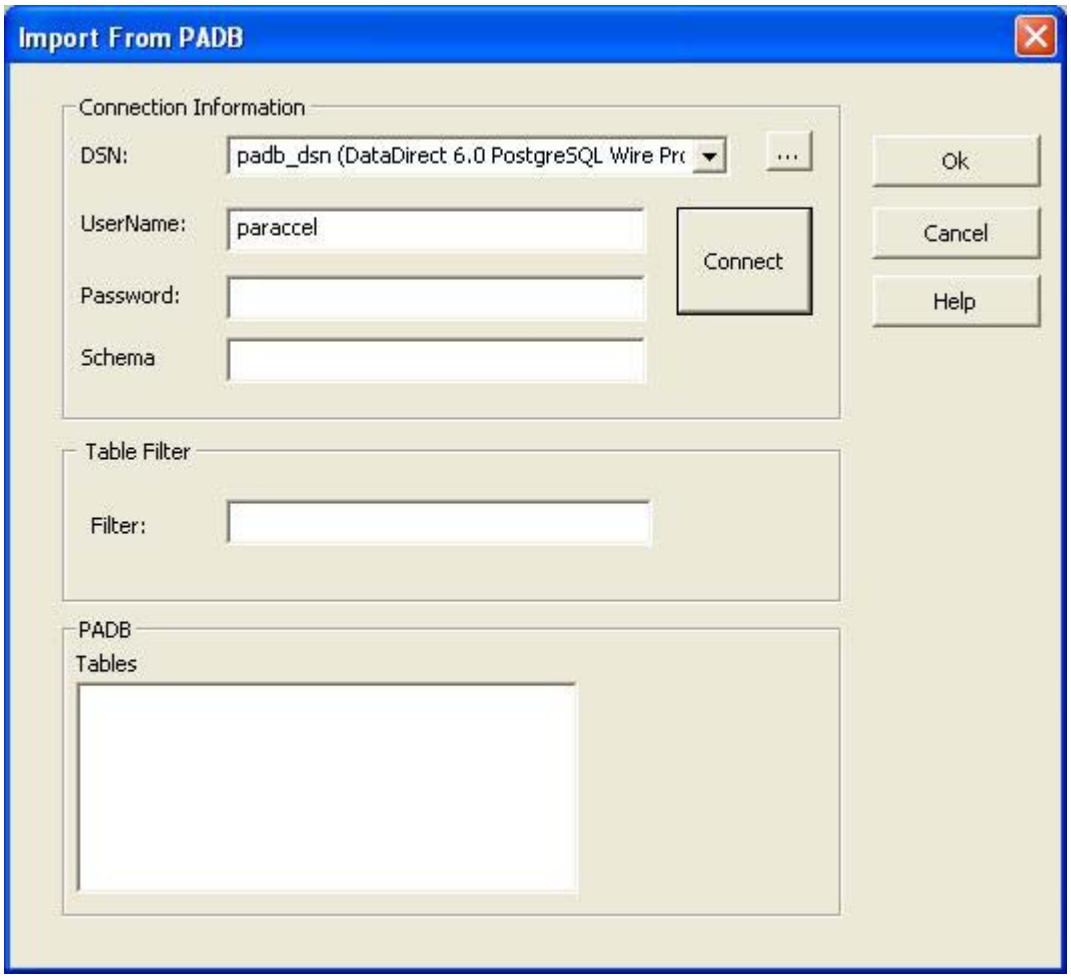
You must specify an ODBC data source in the PowerCenter Designer when you import ParAccel tables as target definitions.

To specify an ODBC data source in the PowerCenter Designer:

1. Connect to the repository from the **Repository Navigator**.

   Note: If you do not connect to a repository, you cannot view the options in the **Tools** menu.

2. Click **Tools** > **Target Designer**.
3. Click **Targets** > **Import from PADB**. The **Import from PADB** dialog box appears.
4. To connect to the ParAccel database, select the ODBC data source that you created in the ODBC Data Source Administrator. Alternatively, you can create or modify an ODBC data source by clicking the ... button.
5. Enter the user name and password to connect to PADB.



Uninstalling the connector on Windows

You can uninstall PowerExchange for ParAccel if you want to remove PowerExchange for ParAccel from Informatica PowerCenter.

If uninstalling the plug-in for the Informatica client, shut down the Informatica client before uninstalling the plug-in. If uninstalling the writer plug-in for the Informatica server, shut down the Informatica server or take the server offline before uninstalling the plug-in.

Note: If you use the plug-in to create repository objects, uninstall the plug-in, and then attempt to open the repository, you might get multiple occurrences of PowerCenter error code REP_57164 with text indicating that an object cannot be loaded because the plug-in used to create it cannot be found. If this happens, click **OK** to clear each message, but when prompted to save your changes, click **No**. If you do not click **No**, your previously created mappings will be overridden.

To uninstall the client or writer plug-in:

1. If uninstalling the client or writer plug-in, use the **Add/Remove Programs** utility in the Windows control panel to uninstall the plug-in.

   If you do not see the plug-in under **Add/Remove Programs**, stop any running PowerCenter client s/w running on the Windows client, and then delete the exchange_padb.chm, pmpadbclient.dll, and padb_EN.res files from the directories listed in *Installing the connector on Windows*.
2. Unregister the writer plug-in in the PowerCenter Administration Console. For more information on unregistering the writer plug-in, see the *PowerCenter Administration Guide*.

Unregistering the repository plug-in

You have to unregister the `PADBPlugin.xml` plug-in from the PowerCenter repository if you want to uninstall PowerExchange for ParAccel. Before unregistering a repository plug-in, ensure that all users are disconnected from the repository and the Repository Service is running in the exclusive mode.

To unregister the repository plug-in:

1. Log into the PowerCenter Administration Console.
2. Change the operating mode of the repository to exclusive.
3. Unregister the `pmpadbwriter.xml` plug-in. For information about uninstalling the plug-in by using the PowerCenter Administration Console, see the *PowerCenter Administrator Guide*.
4. Change the operating mode of the repository to normal.

Note: The list of registered plug-ins for a Repository Service appears in the **Plug-ins** tab of the PowerCenter Administration Console. If the Repository Service is not running in the exclusive mode, the **Remove** buttons for plug-ins already installed are disabled. For more information about exclusive mode, see the *PowerCenter Administrator Guide*.

Unregistering the client plug-in

You have to unregister the `pmpadbclient.dll` file in Microsoft Windows Server 2003 if you want to uninstall PowerExchange for ParAccel.

To unregister the client plug-in:

1. Open the **Registry Editor** in Microsoft Windows Server 2003.
2. Open the following location in the **Registry Editor**:

   ```
   HKEY_LOCAL_MACHINE\SOFTWARE\Informatica\PowerMart
   Client Tools\8.6.1\Plug-ins\Informatica
   ```
3. Delete the string `PC_PWXPADB`.

Caution: Incorrectly editing the registry may severely damage your system. Back up any important data on your computer before making changes to the registry.

Installing the connector on Linux

PowerExchange for PADB allows you to access the PADB metadata stored in the Informatica repository and run sessions against PADB. This section assumes that the Informatica PowerCenter Server is installed on a platform with the 64-bit version of Red Hat Linux 4.0.

You need to install only the writer plug-in, `pmpadbwriter.so`, on Linux Red Hat 4.0 with Informatica server.

Note: The Informatica client resides on a Windows environment machine. The Windows machine contains the files `pmpadbclient.dll` and `PADB_en.res`.

To install PowerExchange for PADB on Linux Red Hat 4.0:

1. Copy the following files to the `<PowerCenter 8.6.1 Installation Directory>/Server/Bin` directory:
   - `pmpadbwriter.so`
   - `padb_EN.res`
2. Copy the `PADBPlugin.xml` file to a temporary location that is accessible from the computer that will run the PowerCenter Administration Console.
3. Register the repository plug-in by using the XML file in the PowerCenter Administration Console on the Windows machine. For more information on registering the repository plug-in, see the *PowerCenter Administration Guide*.
4. Install the PADB ODBC driver in Red Hat Linux 4.0. See *Installing the ODBC driver on Linux*.
5. Set the language code page to `Codepage: UTF-8` in PowerCenter. For more information about setting the language code page, see the *PowerCenter Administrator Guide*.
6. Set the following environment variables:
   - ODBCINI: Point to the `odbc.ini` file in the Informatica server installation directory. For example:

     ```
     $export ODBCINI=
     /home/paraccel/Informatica/32/ODBC6.0/odbc.ini
     ```

   - LD_LIBRARY_PATH: Point to the library path for the ODBC libraries. For example:

     ```
     $export LD_LIBRARY_PATH=
     /home/paraccel/Informatica/32/ODBC6.0/lib:$LD_LIBRARY_PATH
     ```

   - If $LANG is set, unset this environment variable.

Installing the ODBC driver on Linux

PowerCenter uses Open Database Connectivity (ODBC) to connect to ParAccel and import ParAccel metadata. An ODBC data source contains information to access the database. You have to install the PADB ODBC driver after the necessary PowerExchange files are copied and Informatica is configured on the 64-bit Linux system.

PADB supports using PowerCenter8.6.1 installations that have an `ODBC6.0` directory (Hotfixes 11+). If you are using an earlier version of PowerCenter, update your version to 8.6.1, hotfix 11.

To install the ODBC driver on a 64-bit Linux system:

1. Obtain the ODBC RPM files from ParAccel and copy them to a temporary location on the machine where you are installing the driver.

2. Type the following command to install the driver:

   ```
   sudo rpm --force -i driver_name.rpm
   ```

   For example, the following command installs the `padbODBC-08.03.0924-1.x86_64.rpm` driver:

   ```
   $ sudo rpm --force -i padbODBC-08.03.0924-1.x86_64.rpm
   ```

   The `padbodbcw.la` and `padbodbcw.so` driver files are now installed in the `/usr/local/lib64` directory.

For more information, see the PADB *Administrator's Guide and SQL Reference*.

Uninstalling the connector on Linux

You can uninstall PowerExchange for ParAccel from PowerCenter if you want to remove PowerExchange for ParAccel from Informatica PowerCenter.

Note: Do not uninstall the plug-in unless the Informatica server is shut down or has been taken offline.

To uninstall PowerExchange for ParAccel in Linux Red Hat 4.0:

1. Delete the following files from the `<PowerCenter 8.6.1 Installation Directory>\Server\Bin` directory:
   - pmpadbwriter.dll
   - padb_EN.res
2. Unregister the repository plug-in in the PowerCenter Administration Console. For more information, see the *PowerCenter Administration Guide*.

Loading data into PADB tables

The PADB connector plug-in enables you to load data from any source supported in PowerCenter Source Analyzer to the PADB tables. PowerExchange for PADB integrates PowerCenter with PADB by leveraging the COPY command to load data into PADB tables.

Importing source definitions

Source definitions are the metadata of the tables whose data you want to import into PADB. Before creating a source definition, create the source metadata using a wizard or source editor (described in the *PowerCenter Designer Guide*). The PowerCenter repository stores the metadata of both target and source definitions. The metadata describes the schema of the imported definitions. The source can be any file supported in PowerCenter Source Analyzer. Sources can include flat files, XML files, relational tables, COBOL files, and any other type supported by PowerCenter.

To import a source definition:

1. Click **Tools** > **Source Analyzer** in the PowerCenter Designer.
2. Click **Sources** > **Import from File** or **Import from Database** to import source table metadata as a source definition.

Importing target definitions from PADB

Target definitions are the metadata of imported target tables from PADB. The PowerCenter repository stores the metadata of target and source definitions. The metadata describes the schema of the imported definitions.

Before creating a target definition, you must specify an ODBC data source to connect to PADB. From the Control Panel, go to **Administrative Tools** > **Data Sources (ODBC)** to create a data source name (DSN) on the system where the PowerCenter client is installed. Before creating the DSN for PADB, install the PADB ODBC driver. For more information on installing the ODBC driver for PADB, see the PADB *Administrator's Guide and SQL Reference*.

After creating the ODBC data source for PADB, use the Informatica Designer to import the target definition.

To import target definitions from PADB:

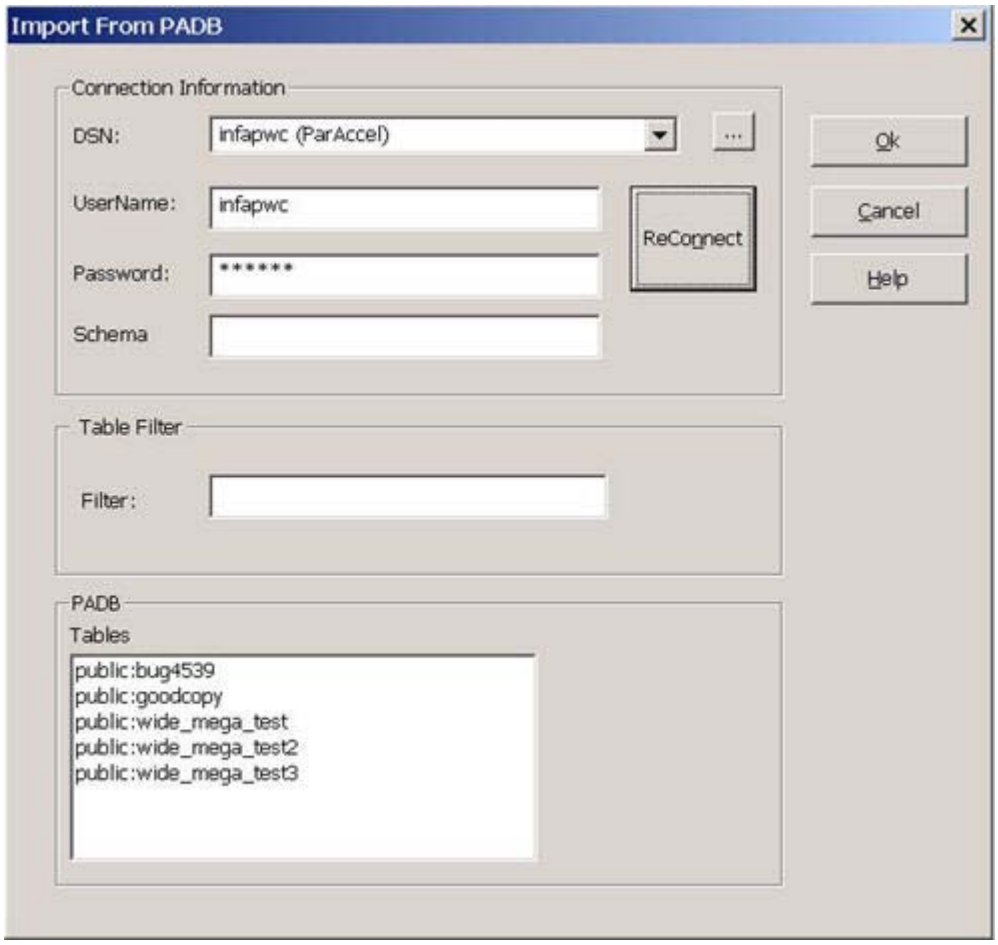
1. Click **Import from PADB** in the Target Designer. The Import from PADB dialog box appears.
2. Select a DSN from the **DSN** drop-down box to specify the data source. (or click to select the DSN).
3. Type your login credentials in the **Username** and **Password** text boxes.
4. Type the filter expression if required. For more information, see *Using filters*.
5. Click **Connect**. The client plug-in displays a list of retrieved tables in the **Tables** field if the connection is successful. The **Connect** button changes to **Reconnect**.
6. Select the required table in the **Tables** list. You can select more than one table using the mouse pointer. To cancel the selected tables, click the selected tables.
7. Click **OK**. The connector plug-in creates the target definitions for the tables selected in the Target Designer.

Note: You can use the **Reconnect** button to refresh the list of tables retrieved as per the filter expression.

Import from PADB dialog box

The **Import from PADB** dialog box enables you to:

- Connect to a single database in PADB (specified by the odbc.ini file)
- Select the DSN
- Import metadata from PADB
- Filter the tables retrieved from PADB



The following table describes the fields in the **Import from PADB** dialog box.

Field/buttons	Description
DSN	Enables you to select the required DSN.
...	Configures the DSN using the ODBC Manager of the system.
Username	Accepts the username required to connect to PADB.
Password	Accepts the password required to connect to PADB.

Field/buttons	Description
Schema	Schema name, if necessary. If this field is blank, the client plug-in retrieves all the table names from all the schemas.
Connect	Connects to the DSN using the credentials given in the User Name and Password text boxes. The connector connects to the PADB leader node. After the connection is successful, the Connect button changes to Reconnect.
Reconnect	Repopulates the tables list.
Filter	Populates the list of tables in the Tables field based on the filter expression given. If no filter is specified, the connector populates the list of all tables in the corresponding schema. You can use special characters `%` and _ in the filter expression. For more information on using regular expressions, see *Using filters*.
PADB Tables	Displays the list of tables that is retrieved from PADB. You can select more than one table. The connector creates the target definitions in PowerCenter Designer for the selected tables when you click OK.
OK	Creates target definitions in PowerCenter Designer for the tables selected in the Tables list and exits the dialog page. **Note**: The client plug-in displays an error message if you do not select any table and click OK.
Cancel	Cancels the operation and exits the dialog box.
Help	Opens the online help for the connector.

Using filters

You can use an expression in the Filter textbox to filter the tables populated in the PADB Tables field. If the Filter text box is left blank or not specified, the connector populates it with all the tables in the corresponding PADB database.

The Filter field accepts basic regular expressions, such as the expressions used with the SQL LIKE predicate. You can use the following characters supported by ODBC in the filter expression:

- `%` (percentage character)

  The percent sign `%` matches zero or more characters.

- _ (underscore character)

  The underscore _ matches a single character.

Examples of filter expressions

Expression	Result
`%ale%`	All tables that have the `ale` substring in their names.
`sale%`	All tables that start with `sale`
`sale_`	All tables that start with `sale` and have only one character after the `e`

PADB data types

Each value that PADB stores has a data type with a fixed set of associated properties. The data types are declared when a table is created, and for stored procedures and functions. These data types constrain the set of values that each column or argument can have.

The following table lists the data types that you can see while importing target definition tables from PADB.

Data Types	Aliases	Description
SMALLINT	INT2	Signed two-byte integer
INTEGER	INT,4	Signed four-byte integer
BIGINT	INT8	Signed eight-byte integer
DECIMAL	NUMERIC	Exact numeric of selectable precision
REAL	FLOAT4	Single precision floating-point number
DOUBLE PRECISION	FLOAT8	Double precision floating-point number
BOOLEAN	BOOL	Logical boolean (true/false)
CHAR	CHARACTER	Fixed-length character string
VARCHAR	CHARACTER VARYING	Variable-length character string with a user-defined limit
DATE		Calendar date (year, month, day)
TIMESTAMP		Date and time (without time zone)

Creating a mapping in PowerCenter

A mapping in PowerCenter enables you to determine the flow of data from sources to targets. You can create and edit a mapping in the Mapping Designer tool.

For more instructions on how to create mappings, see the *PowerCenter Designer Guide*.

After the source and target definitions are created, you have to create a mapping between the source and target definitions.

Creating and running a session

A session enables the Integration Service to move data between sources and targets. After the mapping is created, you have to create a session.

For details on how to create sessions from mappings, see *Sessions* in the *PowerCenter Workflow Basics Guide*.

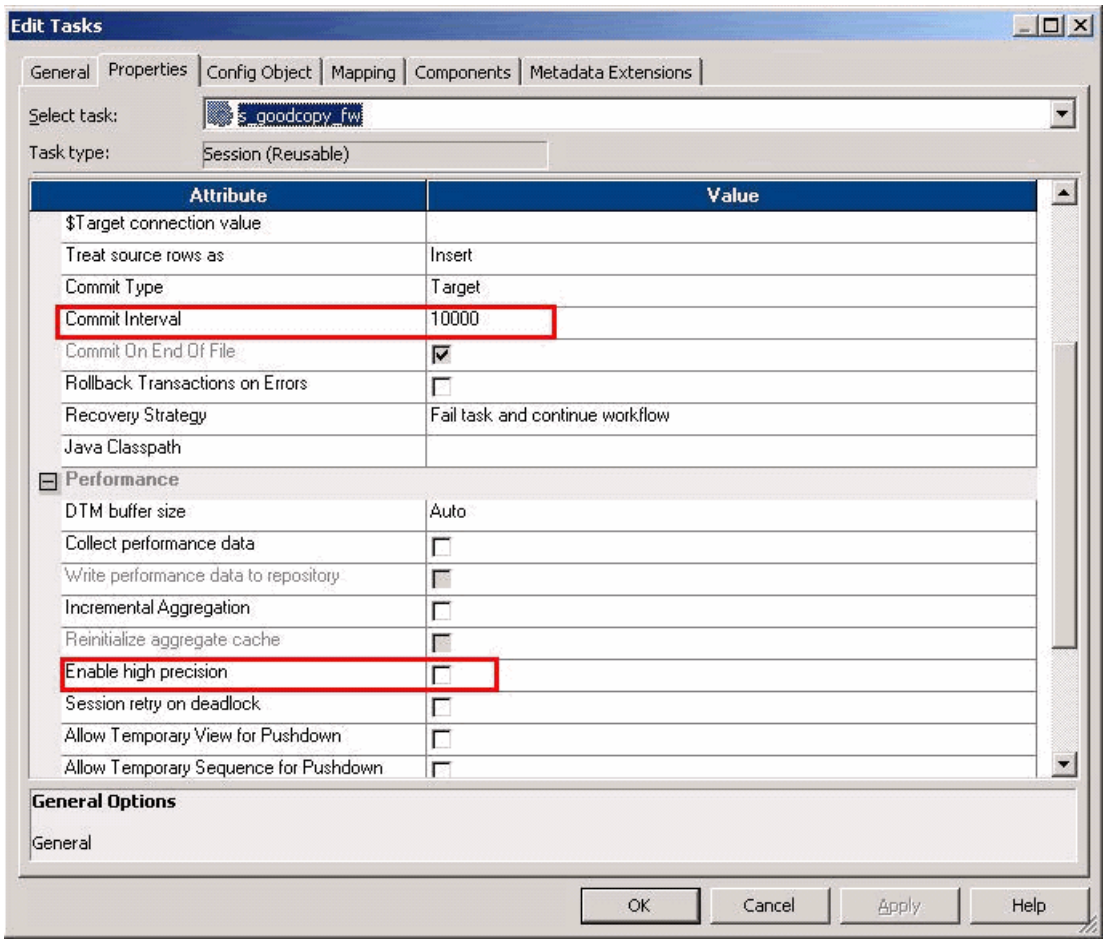
Session-level attributes

Session-level attributes are COPY command options that you can configure. Each session-level attribute is applied individually to each execution of the COPY command.

Some of the session-level attributes are parameterized, which means that the session-level attribute values can be stored in a text file that the Integration Service reads before starting a session. The Integration Service runs the session using the parameter file. The main advantage of using a parameter file is that you can use different parameter files for running the same session based on requirements, and you need not configure the session-level attributes every time you run a session. You can use any text editor to create a parameter file. For more information on using parameter files, see *Parameter Files Overview* in the *PowerCenter Designer Guide*.

Configuring the session Properties tab

The Properties tab contains various general parameters related to a session:



The following table describes the various session Properties attributes that have been customized for PADB or have caveats for use with PADB.

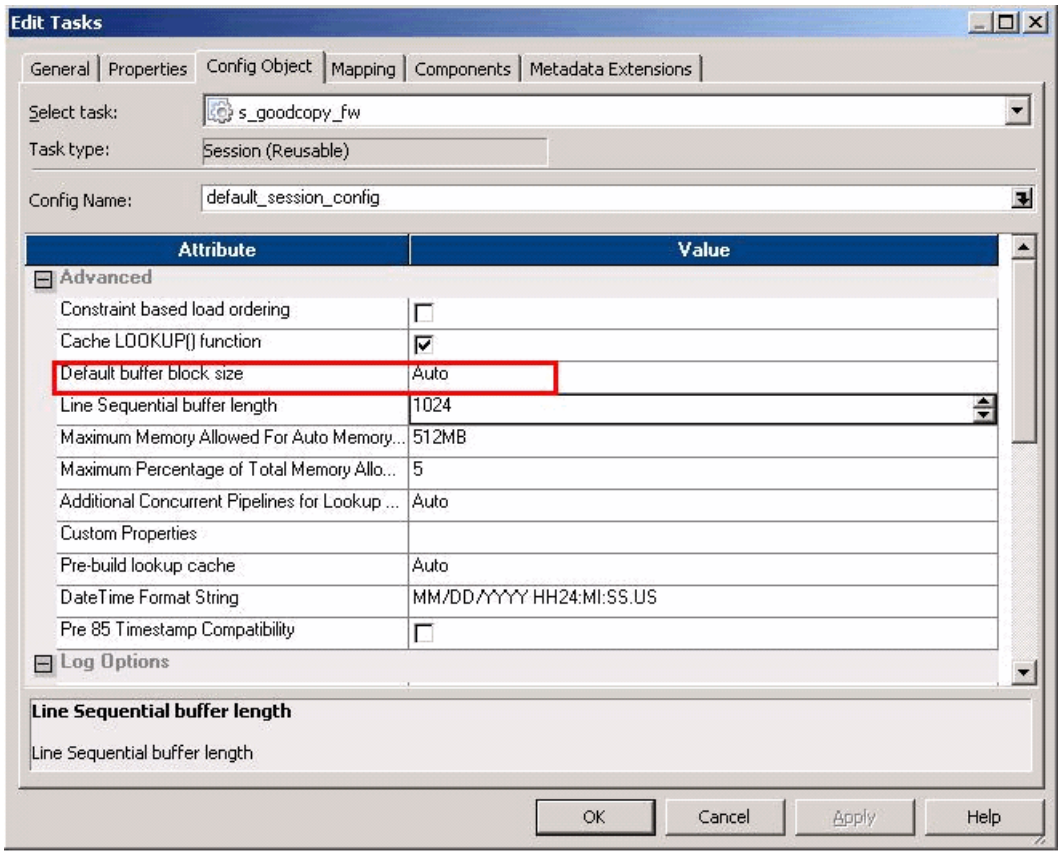
Name	Description
Commit Interval	Uses Informatica's Commit Interval algorithm, not the Commit Interval of PADB's COPY command. Use a large value for large data loads so as not to slow down performance.

Name	Description
Enable high precision	Enable high precision for any data loads containing NUMERIC target columns. If not enabled for numerics, the default disabled value limits precision to 15.

For more information about configuring other Properties attributes, see the *PowerCenter Designer Guide*.

Configuring the Config Object tab

The Config Object tab contains buffer and memory related settings:



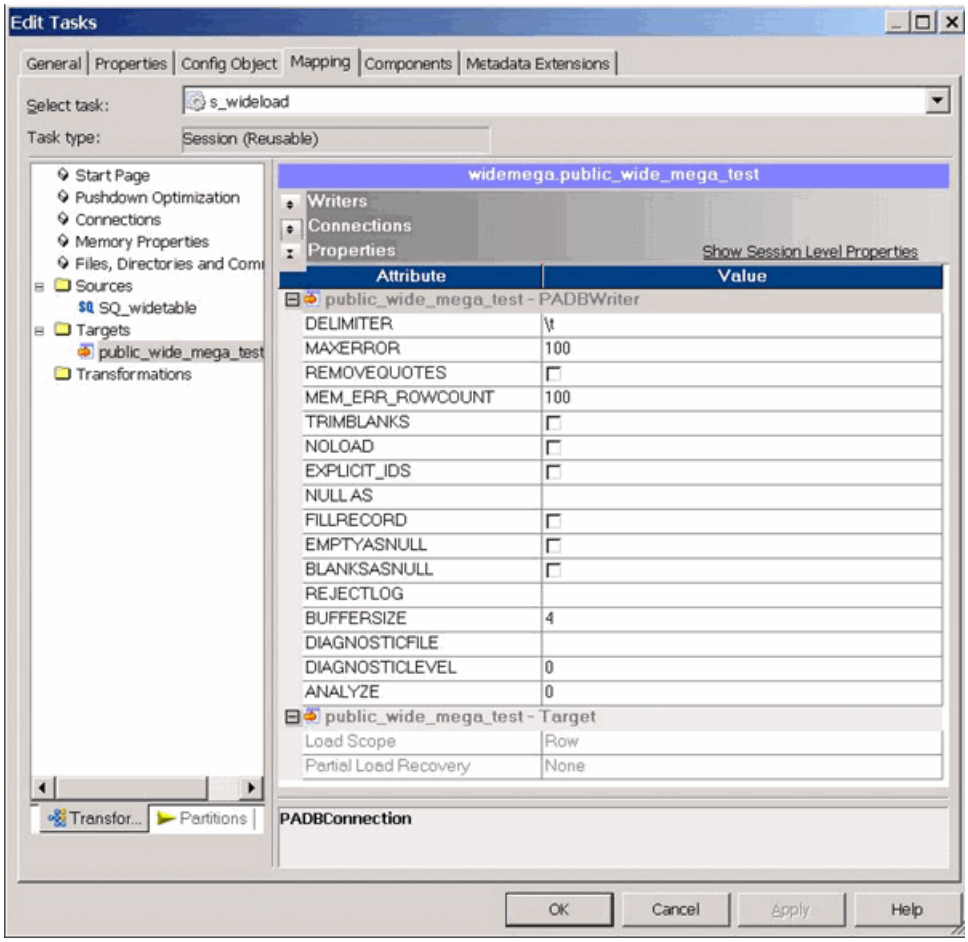
The following table describes the various session Config Object attributes that have been customized for PADB or have caveats for use with PADB.

Name	Description
Default buffer block size	Controls the number of data rows that Informatica supplies to the plug-in at a time. The default value (Auto) for this option supplies about 100 rows. ParAccel recommends setting this value much higher, to at least 10MB.

For more information about configuring other Config Object attributes, see the *PowerCenter Designer Guide*.

Configuring the Mapping tab

The Mapping tab contains attributes related to an already created mapping:



The following table describes the various session Mapping attributes that have been customized for PADB or have caveats for use with PADB. Additional details for the syntax of these commands can be found in the *PADB Administrator's Guide*.

Name	Description
DELIMITER	Single ASCII character that is used to separate fields in the input file, such as a pipe character (|), a comma (,), a tab (\t) or other escape characters, and octals (\011). The default delimiter is a tab character. NULL (\0), LF (\n) and CR (\r) are invalid values. Cannot be used with FIXEDWIDTH. Value should not be surrounded with quotes.
MAXERROR	Maximum number of load errors that can occur before automatically failing the data load. Applies to the COPY command statement, not to the session. The actual number of errors reported might be greater than the specified MAXERROR because of the parallel nature of PADB. If any node in the PADB cluster detects that MAXERROR has been exceeded, each node reports all of the errors it has encountered.
REMOVEQUOTES	Removes surrounding quotation marks from data strings.
MEM_ERR_ROWCOUNT	Sets a limit on the number of error rows to be logged.
TRIMBLANKS	Removes the trailing whitespace characters from a VARCHAR string.
NOLOAD	If enabled, checks the validity of the data file without actually loading the data.
EXPLICIT_IDS	Use EXPLICIT_IDS with tables that have IDENTITY columns to override the auto-generated values with explicit values from the source data files for the tables.

Name	Description
NULL AS	The string that represents a null value. The default is \N (backslash-N).
FILLRECORD	If enabled, allows data files to be loaded when contiguous columns are missing at the end of some of the records.
BLANKSASNULL	If enabled, loads strings that consist of white space characters as NULLs.
REJECTLOG	Specifies a pre-existing directory to store logs for rejected data rows.
BUFFERSIZE	Specifies the size of the buffer to be used by the streaming COPY client to send data to PADB. Adjusting the BUFFERSIZE might improve performance.
DIAGNOSTICFILE (optional)	Full path of a file to receive diagnostic logging from the Informatica plug-in. Defaults to `C:\PADB_INFA_PLUGIN.log` on Windows and `/tmp/PADB_INFA_PLUGIN.log` on Linux.
DIAGNOSTICLEVEL	Specifies the level of detail to be recorded in the diagnostic log. Valid values are 0 (none), 1 (terse), 2 (medium), and 3 (verbose). These diagnostics help with potential debugging issues.
ANALYZE	If set to a value greater than zero, indicates the minimum number of records applied by the load to trigger an ANALYZE *table_name* statement, which will be executed after the COPY command completes. A value of zero disables the automatic ANALYZE.

For more information about configuring other Mapping attributes, see the *PowerCenter Designer Guide*.

Handling errors

By default, the Informatica Integration Service does not stop a session when it encounters non-fatal errors. You can set the error limit for a session by configuring the **Stop on errors** session attribute. The **Stop on errors** attribute is the sum of the PADB errors (MAXERROR) and Informatica PowerCenter related errors. The **Stop on errors** attribute determines the number of non-fatal errors that the Integration Service can encounter before it fails the session. The default value is 0.

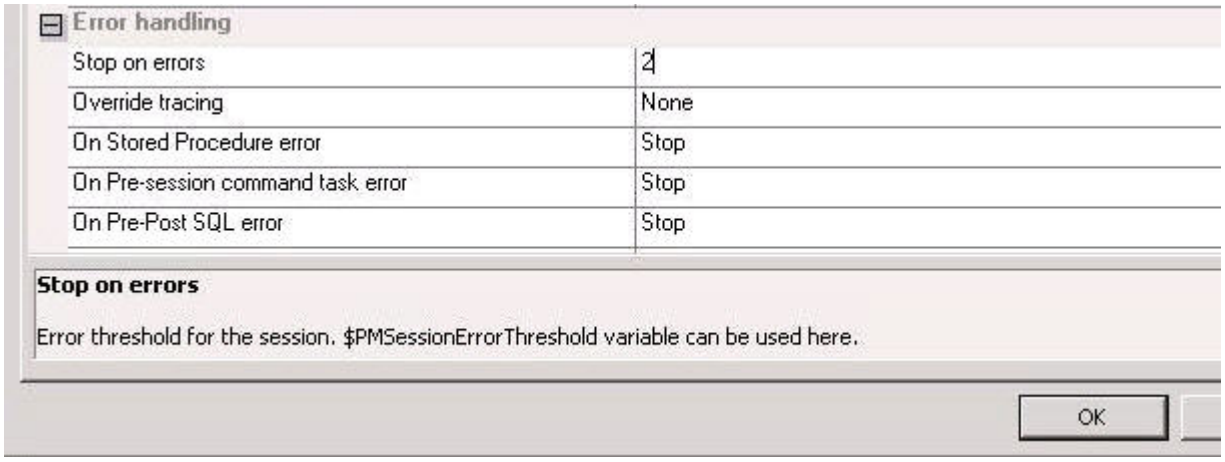
Note: If **Stop on errors** is set to 0, the success or failure of a session depends on the MAXERROR session-level attribute.

The **Stop on errors** limit includes database errors that the COPY command encounters while loading data into the PADB tables. You can also set the error limit for errors encountered during the COPY command execution by configuring the MAXERROR attribute. The Stop on Errors session property and the MAXERROR session property are both used to effectively control sessions.

To specify the Stop on Errors attribute:

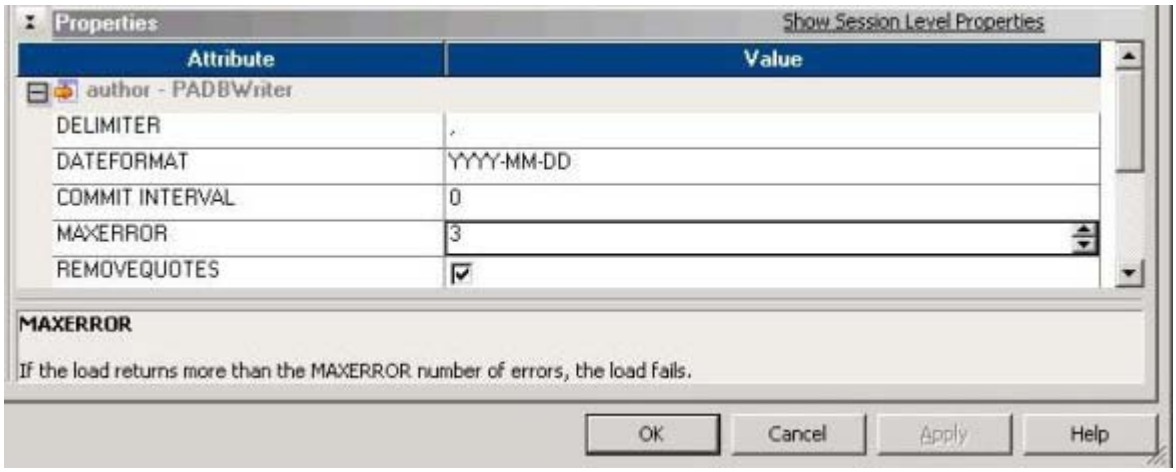
1. Right-click the required session. The **Edit Tasks** dialog box displays.
2. Select the **Config Object** tab.
3. Enter the number of non-fatal errors for the **Stop on errors** attribute.
4. Click **OK**.

Note: The session is always successful if the number of rows is less than the value specified for Stop on Errors.



To specify the MAXERROR attribute:

1. Right-click the required session. The **Edit Tasks** dialog box displays.
2. Select the **Mapping** tab.
3. Enter the number of non-fatal errors for the **MAXERROR** attribute.
4. Click **OK**.



Handling errors: examples

Example 1

- MAXERROR < 3
- Stop on errors < 2
- Errors encountered: MAXERROR = 1, Stop on errors = 1
- Result: The writer plug-in executes the COPY command and the session is successful because both MAXERROR and Stop on Errors are less than the given values.

Example 2

- MAXERROR < 3

- Stop on errors < 6
- Errors encountered: MAXERROR = 4, Stop on errors = 5
- Result: The writer plug-in does not execute the COPY command because MAXERROR is more than the specified value.

Example 3

- MAXERROR < 3
- Stop on errors < 6
- Errors encountered: MAXERROR = 2, Stop on errors = 6
- Result: The writer plug-in executes the COPY command. The session is not successful because the Stop on Errors condition is not satisfied.

Creating a workflow

The workflow allows the Integration Service to execute a session. After you create a session, you have to create a corresponding workflow.

Note: When Informatica loads data into a target table from a source file, it transforms zero length strings (empty fields) and null values as null in the target table; however, if Informatica is loading data into a target table from a relational database source, it does not transform zero-length strings or null values.

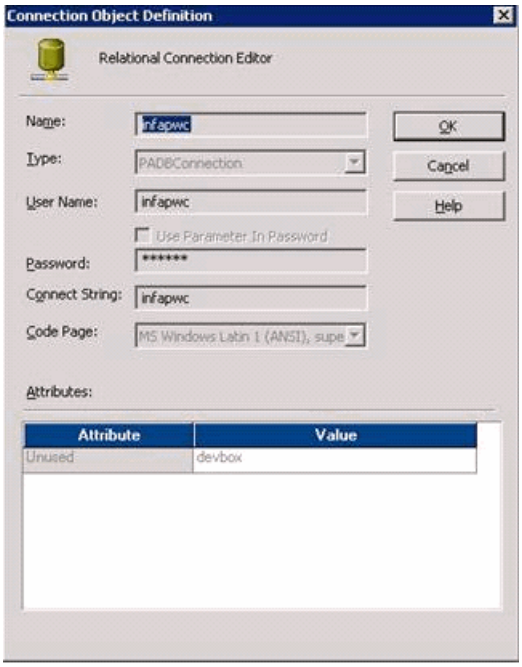
For information about creating and running workflows, see the *PowerCenter Workflow Basics Guide*.

Configuring a relational connection for PADB

You have to configure a relational connection for PADB before you can connect to a PADB database.

To create a new relational connection:

1. Click **Connection** > **Relational Connection**.
2. Click **New** in the Relational Connection Browser window.
3. Enter values for all of the fields and attributes in the Connection Object Definition window:



4. Click **OK**.

The following table describes each field.

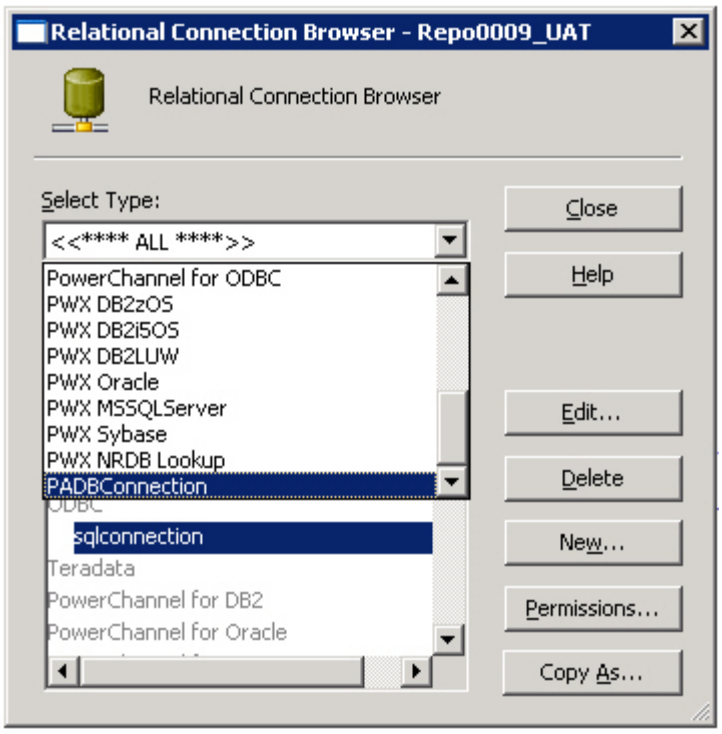
Name	Description
Name	Name of the relational connection
Type	PADBConnection (should be pre-selected)
User Name	Database user name to connect to PADB.
Password	Password used in authentication while connecting to PADB.
Connect String	The connection string (DSN name) that is required to connect to PADB.

For more information about configuring relational connections, see the *PowerCenter Designer Guide*.

Selecting a relational connection for PADB

To select a relational connection for PADB:

1. Click **Connections** > **Relational** in Informatica PowerCenter Manager. The Relational Connection Browser opens.



2. Select **PADBConnection** from the **Select Type** drop-down list, which contains all of the available connection objects.
3. Select the required PADB connection object from the **Objects** list.
4. Click **Edit** if you want to modify the attributes of the relational connection object.
5. Click **Close**.

Error messages

PowerExchange for ParAccel generates error messages that help you troubleshoot loads into PADB. The error messages are classified as follows:

- Client-related error messages. These messages are related to the client plug-in.
- Writer-related error messages. The writer plug-in writes error messages to the Informatica session log.
- PADB also generates error messages, and the writer plug-in retrieves these error messages from the following PADB system tables:
  - STL_LOAD_ERRORS
  - STL_QUERYTXT
  - STL_QUERY

Client-related error messages

Client-related error messages are generated by the client plug-in while connecting to PADB or retrieving target definitions from PADB.

Error messages	Cause/solution
The User is not associated with a trusted PADB server connection.	Check the login credentials for the PADB ODBC DSN connection.
Login Failed for User.	Check the login credentials for the PADB ODBC DSN connection.
Select item(s) you want to import.	You will see this error if you click **OK** without clicking **Connect**. Click **Connect** and then click **OK**.
Nothing has been selected to import.	Select one or more tables in the **PADB Tables** field and click **OK**.
Unable to Import the target table.	Check with the PADB administrator to make sure that the table exists.

Writer-related error messages

Writer related error messages are messages generated by the writer plug-in while connecting to PADB and loading data into the PADB tables.

PADB also generates error messages that the writer plug-in retrieves from the PADB error tables. The writer updates the Informatica session log with these error messages.

Note: If the error message generated is not loaded in any of the PADB system tables, the writer plug-in logs the following error message in the Informatica session log:

```
Error failed to load into stl_load_errors
```

Error messages	Solution
An unknown exception occurred while initializing the plug-in.	Check that the correct version of the writer plug-in is deployed in the server bin directory.

Error messages	Solution
An unknown exception occurred while de-initializing the plug-in.	Restart Informatica.
Failed to retrieve the target list.	Restart Informatica.
Failed to retrieve the target driver.	Restart Informatica.
Failed to retrieve the group list.	Restart Informatica.
Failed to retrieve the group driver.	Restart Informatica.
Failed to retrieve the partition driver.	Restart Informatica.
Invalid session extension.	Restart Informatica.
Failed to allocate memory for the target driver.	Check Informatica prerequisites.
Failed to allocate memory for the group driver.	Check Informatica prerequisites.
Failed to retrieve the partition driver.	Restart Informatica.
Failed to allocate memory for the partition driver.	Check Informatica prerequisites.

The following table shows writer-related error messages retrieved from the PADB tables.

Error messages	Solution
Failed to retrieve session attributes.	Check session attributes.
The session failed because the delimiter specified in the Delimiter session property is invalid.	Check the DELIMITER session attribute.
Invalid Null AS value	Check the NULL AS session attribute.
The session failed because the max error specified in the Max Error session property is invalid.	Check the MAXERROR session attribute.
The Integration Service failed to retrieve the value for the session attribute *session_attribute*.	Check the value for the corresponding session attribute.
Error failed to load into stl_load_errors.	Contact the PADB administrator to find the reason for the error.
Number of rejected rows meets the stop on errors limit.	The session has failed. Check the value for the Stop on errors attribute.
Failed to retrieve the connection attributes.	Check the connection attributes.
The table does not exist in the connected database.	Check if the table exists in the PADB database.

Error messages	Solution
Failed to extract the value from the parameterized attribute.	Check the parameter name.

Messages in Informatica session log

The writer plug-in updates the Informatica session log with error messages. The messages are classified into various severity levels:

Severity level	Description
FATAL	Fatal error.
ERROR	Service failed to perform an operation or respond to a request from a client application.
WARNING	Service is performing an operation that may cause an error. This can cause repository inconsistencies.
INFO	Service is performing an operation that does not indicate errors or problems.
TRACE	Message size.
DEBUG	Success or failure of service operations.

The session log in Informatica also contains Target Table Load Summary information, which shows the following details:

Detail	Description
Requested	The number of rows requested by PowerCenter, or the number of rows that the writer plug-in reads from the PowerCenter memory buffer to load into the PADB target table.
Applied	The number of rows sent to the COPY command by the writer plug-in.
Rejected	The number of rows rejected by the COPY command.
Affected	The number of rows inserted into PADB target tables.

Glossary and abbreviations

Exclusive mode

In exclusive mode, the Informatica repository enables the user to perform administrative tasks. Only one user can access the repository when the repository runs in exclusive mode.

Mapping

A mapping in PowerCenter describes the flow of data from sources to targets.

PowerCenter

PowerCenter is used for accessing and integrating data from various business systems.

PowerCenter buffers

The PowerCenter Integration Service allocates memory blocks while initializing a session. The memory blocks are allocated individually to hold both source data and target data. The Integration Service fails the session if not enough memory is available.

Plug-in

ParAccel provides two plug-ins to use PowerExchange with a PADB database. The available plug-ins are a writer plug-in for a PowerExchange server or a client plug-in for a PowerExchange client.

PowerExchange

PowerExchange is a data access product that you can use to extract, transform, and load data between various enterprise data formats.

Session

A session encloses a single mapping, enabling the Integration Service to move data between sources and targets.

Workflow

A workflow allows the Integration Service to execute one or more sessions.

The following is a list of acronyms and abbreviations used in this guide.

Abbreviation	Full name
PADB	ParAccel Analytic Database
ODBC	Open Database Connectivity
DSN	Database Source Name
DBMS	Database Management System
ETL	Extract, Transform, and Load
MPP	Massively Parallel Processing
DLL	Dynamic Linked Library
OS	Operating System